

Outcomes of Patients with Differentiated Thyroid Carcinoma Following Initial Therapy*

Jacqueline Jonklaas,¹ Nicholas J. Sarlis,^{2‡} Danielle Litofsky,² Kenneth B. Ain,³ S. Thomas Bigos,⁴ James D. Brierley,⁵ David S. Cooper,⁶ Bryan R. Haugen,⁷ Paul W. Ladenson,⁸ James Magner,⁹ Jacob Robbins,¹⁰ Douglas S. Ross,¹¹ Monica Skarulis,¹⁰ Harry R. Maxon,¹² and Steven I. Sherman²

This analysis was performed to determine the effect of initial therapy on the outcomes of thyroid cancer patients. The study setting was a prospectively followed multi-institutional registry. Patients were stratified as low risk (stages I and II) or high risk (stages III and IV). Treatments employed included near-total thyroidectomy, administration of radioactive iodine, and thyroid hormone suppression therapy. Outcome measures were overall survival, disease-specific survival, and disease-free survival. Near-total thyroidectomy, radioactive iodine, and aggressive thyroid hormone suppression therapy were each independently associated with longer overall survival in high-risk patients. Near-total thyroidectomy followed by radioactive iodine therapy, and moderate thyroid hormone suppression therapy, both predicted improved overall survival in stage II patients. No treatment modality, including lack of radioactive iodine, was associated with altered survival in stage I patients. Based on our overall survival data, we confirm that near-total thyroidectomy is indicated in high-risk patients. We also conclude that radioactive iodine therapy is beneficial for stage II, III, and IV patients. Importantly, we show for the first time that superior outcomes are associated with aggressive thyroid hormone suppression therapy in high-risk patients, but are achieved with modest suppression in stage II patients. We were unable to show any impact, positive or negative, of specific therapies in stage I patients.

Introduction

THE INCIDENCE OF epithelial-derived, differentiated thyroid cancer (DTC) has been steadily increasing for at least two decades (1). Despite excellent outcomes, with a lower risk for death or disease recurrence than patients with most other malignancies, mortality rates may be increasing in some subgroups (2). However, adverse events often occur years or decades after diagnosis. Randomized trials of treatment for DTC have, therefore, not been performed, as they would require extraordinarily large numbers of patients to be followed for extended periods of time. In the absence of such trials, passionately defended schools of thought have

evolved regarding the benefit contributed by total or partial thyroidectomy, post-operative radioactive iodine-131 (RAI), and thyroid hormone suppression therapy (THST). In addition, the risks of death and disease recurrence for individual patients are imperfectly predicted by even the best-performing staging systems (3). Lack of randomized trials and the limitation of staging systems both compound the dilemma for clinicians deciding among therapeutic options.

The choice of initial surgery for DTC is influenced by the additional risks of total or near-total thyroidectomy (NTT), compared with lesser surgical resection (4,5). These increased surgical risks are weighed against potential benefits, including the ability to employ more accurate surveillance after

¹Georgetown University Medical Center, Washington, D.C.

²University of Texas M.D. Anderson Cancer Center, Houston, Texas.

³University of Kentucky Medical Center, Lexington, Kentucky.

⁴Maine Medical Center, Portland, Maine.

⁵Princess Margaret Hospital, Toronto, Ontario, Canada.

⁶Sinai Hospital, Baltimore, Maryland.

⁷University of Colorado at Denver and Health Sciences Center, Aurora, Colorado.

⁸Johns Hopkins University School of Medicine, Baltimore, Maryland.

⁹Genzyme Corporation, Cambridge, Massachusetts.

¹⁰National Institutes of Health, Bethesda, Maryland.

¹¹Massachusetts General Hospital, Boston, Massachusetts.

¹²University of Cincinnati Medical Center, Cincinnati, Ohio.

*Presented in part at the 74th Annual Meeting (2001) and the 2nd Annual Spring Meeting (2005) of the American Thyroid Association.

‡Present address: Sanofi-Aventis Pharmaceuticals, Bridgewater, New Jersey.

NTT (6). In some studies, NTT has been associated with improved survival and decreased recurrence (7–9), although other studies suggest that the benefit is limited to high-risk patients (10,11) or that there is no benefit at all (12,13).

Similar uncertainty exists regarding the indications for initial RAI therapy. Although lower recurrence rates (7,14–18) and improved survival (14,17–21) have been reported after RAI, some reports show a survival benefit only in high-risk patients (10,22–25). In one large cohort of DTC patients, RAI did not decrease mortality or recurrence rates in low-risk patients (26). Instead, a trend towards increased recurrence was associated with RAI, potentially reflecting the ability to perform more accurate surveillance for recurrence after ablation of postoperative thyroid remnants. Overall, a recent meta-analysis concluded that definitive data regarding the benefit of postoperative RAI was lacking for low-risk patients (24).

Controversy also exists regarding the extent of benefit from, and optimal intensity of, THST. Such therapy theoretically reduces the ability of thyroid-stimulating hormone (TSH) to act as a growth factor for DTC (27,28). In several series reporting benefit from thyroid hormone therapy (11,15,29–38), lack of thyroid hormone use may have indicated that patients had undergone less extensive surgery. Thus, patients potentially at greater risk for recurrence would be automatically assigned to the group that did not receive thyroid hormone replacement. Only one retrospective study has reported that an undetectable TSH value was associated with longer relapse-free survival, independent of the initial disease stage (39). A meta-analysis that quantified adverse clinical outcomes from 10 studies of THST also suggested a benefit of this management approach (40). The authors of a recent review of THST concluded that there was limited evidence of benefit of THST, except in patients with persistent or recurrent disease (41).

The National Thyroid Cancer Treatment Cooperative Study Group maintains a registry currently contributed to by 11 North American institutions. The registry prospectively follows a large nonrandomized cohort of patients with DTC, with the objective of assessing the effects of initial and longitudinal management on their outcomes. Previously, the registry investigators have described a unique staging system with excellent prognostic capabilities (3), reported the

survival benefit associated with NTT and RAI in 385 stage III and IV patients (22), and identified a tendency towards less disease progression associated with THST in 141 patients with stage III and IV papillary thyroid carcinoma (38). We now report the findings from our larger analysis of 2936 DTC patients of all stages, examining the relationship between initial treatment and patient outcomes.

Methods

Patients and data collection

The data collection and analysis methods of the registry have been described previously (3,22,38). New patients are registered on an ongoing basis within 3 months of their initial surgery. Between January 1987 and June 2001, 2936 patients were registered. Patients registered at each institution received the treatment and follow-up that the managing physician determined to be appropriate, independent of registry participation. Surgery, RAI treatment administered within 6 months of the patient's surgery, and THST were considered to be initial therapies. The extent of thyroid surgery was assessed by each institutional investigator. RAI was classified as "adjuvant" if the patient had no evidence of residual disease at the time of its administration. All variants of papillary thyroid cancer (PTC) were included together, whereas the follicular thyroid cancer (FTC) group included the Hürthle cell variant (HCC). Disease stage was assigned using the previously described registry staging system (3) (Table 1) and the AJCC-TNM5 system (42). (The TNM5 staging is retained both because it allows comparison with previous studies examining DTC outcomes and because our previous data collection parameters did not permit classification according to the revised AJCC-TMN6 staging criteria.) Clinical status at entry was classified by the individual investigator at each participating institution as no residual disease, or residual disease with documentation of the extent of remaining tumor or metastases. Patients initially assigned to stages I and II and to stages III and IV were designated as "low-risk" and "high-risk" patients, respectively.

Follow-up data were prospectively collected for each patient on an annual basis, as described previously (3,22,38), and data through 2001 were included in this analysis. Extent of disease and disease progression were assessed by in-

TABLE 1. REGISTRY STAGING CLASSIFICATION

		Papillary carcinoma		Follicular carcinoma	
		Age < 45	Age ≥ 45	Age < 45	Age ≥ 45
Primary Tumor Size (cm)	< 1	I	I	I	II
	1–4	I	II	I	III
	> 4	II	III	II	III
Primary Tumor Description	Microscopic multifocal	I	II	I	III
	Macroscopic multifocal or macroscopic tumor capsule invasion	I	II	II	III
	Microscopic extraglandular invasion	I	II	I	III
	Macroscopic extraglandular invasion	II	III	II	III
	Poor differentiation	n/a	n/a	III	III
Metastases	Cervical lymph node metastases	I	III	I	III
	Extracervical metastases	III	IV	III	IV

dividual investigators. When possible the causes of death were reviewed and mortality data were confirmed through the Social Security Death Index. Recorded data were entered into a PC-based clinical data management system (Medlog®, version 4-2000, Incline Village, NV) at each site, and transmitted to a central data repository, currently maintained at The University of Texas M.D. Anderson Cancer Center. Approval was obtained from the Institutional Review Board (IRB) at each participating institution, and informed consent obtained from patients prior to registration, if required. IRB approval was also obtained at M.D. Anderson Cancer Center for maintaining the central data repository and performing subsequent analyses. Collected data were used to determine the overall survival (OS), disease-specific survival (DSS), and disease-free survival (DFS) of registry participants. These parameters were examined according to disease stage.

TSH scores

TSH scores were generated for each patient as previously described (38). During the data collection period, second-generation or third-generation TSH assays were in use at each institution's clinical laboratory, with a functional sensitivity of at least 0.1 mIU/L. Based upon the functional sensitivities and normal ranges in use at each institution, each serum TSH level obtained was characterized as "undetectable," "subnormal," "normal," or "elevated." A mean TSH score was calculated for each patient from all available TSH results using the following ordinal scale: 1 = undetectable, 2 = subnormal, 3 = normal, 4 = elevated. Subsequently, patients were divided into three groups based on their mean TSH scores: group 1 (mean TSH scores 1.0–1.9), group 2 (mean TSH scores 2.0–2.9), and group 3 (mean TSH scores 3.0–4.0).

Statistical analyses

Chi-square analysis and *t*-tests were performed for nominal and normally distributed continuous data, respectively. Product-limit survival analysis (using log-rank statistic) was performed to determine the univariate predictors of survival. The specific therapies examined in our analysis were a less extensive thyroid surgery versus NTT, no RAI versus post-operative RAI (given for any reason or as adjuvant therapy), and increasing degrees of THST. The outcomes analyzed were OS, DSS, and DFS. The outcomes of all patients with follow-up were considered in the analysis of OS and DSS, whereas only those patients thought to have no residual disease after initial therapy were analyzed for DFS. Proportional hazards multivariate regression modeling was performed to determine the relative contribution of each predictor to OS, DSS, and DFS. Risk ratios were calculated from the proportional hazards analyses along with the corresponding 95% confidence intervals. A *p* value was also generated for the "best model" which incorporated the set of therapies that best predicted outcome for a particular group of patients. Throughout this report, a risk ratio for two therapies denoted as "A/B" that is less than one indicates a worse outcome associated with therapy B compared with therapy A, whereas a risk ratio greater than one indicates an improved outcome associated with therapy B.

Propensity analysis (43,44) was performed to evaluate the role of potential selection bias in the application of therapies. By identifying cohorts of patients with similar *a priori* prob-

ability or "propensity" to receive a particular treatment based on their clinical characteristics, propensity analysis can remove up to 90% of the bias inherent in comparing treatments used in a nonrandomized cohort study. The probability that a patient would receive a particular treatment was calculated with a stepwise regression model that incorporated gender, age, histology, tumor size, extraglandular invasion, neck metastases, surgical extent, reporting institution, and year of diagnosis. The goodness of fit of the propensity score, determined by the area under the receiver operator characteristic curve for this regression model, was calculated to determine if the model discriminated well between patients who received the treatment and those who did not, thus balancing covariates that might have influenced outcomes. Patients were then separated into strata by propensity score rank. Within each propensity score stratum of patients with similar likelihood of having received a treatment, the difference in the outcome between those receiving and not receiving the treatment could be examined.

A *p* value of 0.05 or less was considered statistically significant. All analyses were performed using the SAS JMP 5 (version 5.0.1.2, copyright 1989–2003) statistical software package available at the coordinating center.

Results

Characteristics of the cohorts

Between January 1987 and June 2001, 2936 patients with DTC were registered from the 11 currently participating institutions. None of the participating institutions contributed more than 15% of the patients in the database. Eighty-six percent had a diagnosis of PTC, 10% had FTC, and 4% had HCC. Median duration of follow-up was 3 years (range 0–14), for a total of 10,994 person-years of documented follow-up. The average number of TSH values per patient in the entire cohort was 2.3. The characteristics of this overall cohort are shown in Table 2a. Seventy percent and 73% of the patients were considered low risk by registry and AJCC-TNM5 staging respectively. There were 152 of the 2936 patients with zero follow-up time. These patients were not included in survival analyses. During follow-up, 153 patients died (5% of the overall cohort), with 81 disease-specific deaths (79 attributed to DTC and 2 from complications of therapy). Of the 2204 patients considered disease-free after primary therapy, 221 (10%) were diagnosed with recurrent disease within a median of 1.5 years from diagnosis. Distant metastases accounted for 18% of the recurrences, compared with 71% for local or regional recurrence, and 11% for whom site of recurrence was not reported. Five-year overall survival following recurrence was 87% for local or regional disease, compared with 72% after distant metastasis.

Similar to our previous report (38), impact of THST was analyzed in a subgroup of patients in whom 50% or more of their TSH values were recorded, subsequently referred to as the "THST cohort." The average number of TSH values per patient over the study period was 3.9 (range 1–12). The characteristics of the 1548 patients in the THST cohort are also shown in Table 2a. There were no significant differences between the overall and THST cohorts for these clinical characteristics. The characteristics of the cohort remaining after the THST cohort had been extracted from the overall cohort are shown in Table 2b. This "remaining" cohort had

TABLE 2a. CLINICAL CHARACTERISTICS OF ANALYZED COHORTS

Parameter		Overall and Disease-Specific Survival		Disease-Free Survival	
		Overall Cohort	THST Cohort ^b	Overall Cohort	THST Cohort ^b
Number of Patients		2936	1548	2204	1209
Gender	M	28%	26%	26%	25%
	F	72%	74%	74%	75%
Age at Dx	< 20	4%	3%	3%	2%
	20–45	53%	54%	54%	55%
	> 45	43%	43%	42%	43%
Histology	Papillary	86%	86%	87%	87%
	Follicular	10%	9%	9%	9%
	Hürthle	4%	4%	4%	4%
Registry / TNM5 Stage	I	44% / 56%	44% / 57%	49% / 60%	48% / 60%
	II	26% / 17%	27% / 17%	28% / 18%	29% / 18%
	III	25% / 22%	25% / 22%	22% / 21%	23% / 22%
	IV	5% / 5%	4% / 4%	1% / 1%	< 1% / < 1%
Surgical Treatment	NTT	85%	86%	86%	86%
	Other	15%	14%	14%	14%
RAI therapy— Administered Activity	Yes	70%	75%	68%	73%
	≤ 30 mCi	10%	12%	12%	14%
	31–75 mCi	5%	7%	6%	8%
	> 75 mCi	55%	56%	50%	50%
	No	30%	25%	32%	27%
Adjuvant RAI ^a — Administered Activity	Yes			68%	73%
	≤ 30 mCi			12%	14%
	31–75 mCi			6%	8%
	> 75 mCi			50%	50%
	No			32%	27%
TSH score category ^c	1.0–1.9		43%		43%
	2.0–2.9		46%		47%
	3.0–4.0		11%		10%
Number of deaths		153	59	67	30
Number of disease-related deaths		81	29	30	14
Number of recurrences				221	146

^aentry status: disease-free.

^bsubset of patients who had at least 50% of TSH values recorded.

^cTSH score category: mean of TSH score over all recorded values, where score 1 = undetectable, 2 = subnormal, 3 = normal, 4 = elevated, THST = thyroid hormone suppression therapy.

more patients under 20 and over 45 years of age than the THST cohort, was less likely to have received RAI therapy, and had more stage IV patients. The clinical characteristics of the overall cohort and the remaining cohort overlapped and they were without statistical differences.

Registry disease stage

Stage I patients were 92.4% PTC, 6% FTC, and 1.6% HCC. As registry stage increased the percentage of PTC patients decreased, but the percentage of FTC and HCC patients increased (see Fig. 1a). Figures 1b and 1c show OS and DSS for the overall cohort divided according to the four disease stages. Outcomes were similarly distributed when evaluated by AJCC-TNM5 stage (data not shown). Consistent with our earlier report, initial disease stage was a significant predictor of both OS and DSS ($p < 0.0001$). DFS was examined in patients who had no evidence of disease after their initial therapy. Excluding patients with initial extracervical metas-

tases, disease stage was also predictive of DFS, with 10-year risks of recurrence of 10%, 15%, and 25% for stages I, II, and III, respectively (Fig. 1d).

OS and DSS

OS (Fig. 1e) and DSS (data not shown) were significantly worse for patients with FTC or HCC, compared with PTC. However, analyses of the outcomes associated with various therapies demonstrated no significant differences among the histologic subtypes (data not shown), and therefore these histologies were combined for subsequent analyses. Similarly, patients in stages III and IV were combined, as treatment effects did not differ significantly between these two high-risk groups due to smaller sample sizes. The OS of the THST cohort was significantly longer than that of the remaining cohort (Fig. 1f). OS of the overall cohort and remaining cohort overlapped with each other, resulting in no statistical difference.

TABLE 2b. COMPARISON OF CLINICAL CHARACTERISTICS OF THST AND REMAINING COHORTS

Parameter		Overall and Disease-Specific Survival		p-value ^d	Disease-Free Survival ¹		p-value ^d
		THST Cohort ^b	Remaining Cohort		THST Cohort ^b	Remaining Cohort	
Number of Patients		1548	1388		1209	995	
Gender	M	26%	29%	n.s.	25%	28%	n.s.
	F	74%	71%		75%	72%	
Age of Dx	<20	3%	5%	0.02	2%	5%	0.0036
	20–45	54%	51%		55%	54%	
	>45	43%	44%		43%	42%	
Histology	Papillary	86%	86%	n.s.	87%	88%	n.s.
	Follicular	9%	10%		9%	8%	
	Hürthle	4%	4%		4%	4%	
Registry/TNM5 Stage	I	44%/57%	44%/55%	(NTCTCS) p = 0.026 (TNM5)	48%/60%	50%/61%	(NTCTCS) n.s. (TNM5) n.s.
	II	27%/17%	24%/17%		29%/18%	28%/18%	
	III	25%/22%	25%/21%		23%/22%	21%/20%	
	IV	4%/4%	7%/6%		<1%/<1%	1%/<1%	
Surgical Treatment	NTT	86%	84%	n.s.	86%	86%	n.s.
	Other	14%	16%		14%	14%	
RAI therapy—Administered Activity	Yes	75%	65%	<0.0001	73%	62%	<0.0001
	≤30 mCi	12%			14%		
	31–75 mCi	7%			8%		
	>75 mCi	56%			50%		
Adjuvant RAI ^a —Administered Activity	No	25%	35%		27%	38%	
	Yes				73%	62%	
	≤30 mCi				14%		
	31–75 mCi				8%		<0.0001
>75 mCi				50%			
TSH score category ^c	No				27%	38%	
	1.0–1.9	43%			43%		
	2.0–2.9	46%			47%		
Number of deaths	3.0–4.0	11%			10%		
		59	94	0.0003	30	37	n.s.
Number of disease-related deaths		29	52	0.0019	14	16	n.s.
Number of recurrences					146	75	0.0003

^aentry status: disease-free.

^bsubset of patients who had at least 50% of TSH values recorded, THST = thyroid hormone suppression therapy.

^cTSH score category: mean of TSH score over all recorded values, where score 1 = undetectable, 2 = subnormal, 3 = normal, 4 = elevated.

^dChi-square, for difference between THST cohort and remaining cohort.

¹N = 106 excluded because status at entry undeterminable.

Effect of extent of surgery in the overall cohort

The relationship between extent of surgery and subsequent outcomes is shown in Table 3. By univariate analysis of the overall cohort, NTT was associated with significantly improved OS in stage II and high-risk patients, as compared with a less complete surgery. However, the extent of thyroid surgery was not associated with altered OS in stage I patients. There was no significant association of extent of surgery with DSS or DFS for any stage. Propensity analysis did not alter any of the conclusions regarding the impact of surgery.

Effect of RAI therapy in the overall cohort

Postoperative RAI therapy, administered as adjuvant therapy or to treat residual disease, was associated with

significant differences for all outcomes being evaluated (Table 4). By univariate analysis, any postoperative administration of RAI was associated with significantly improved OS in stage II and high-risk patients, compared with the absence of RAI. DSS was improved in high-risk patients treated with RAI, but not in stage II patients. Adjuvant RAI, given to high-risk patients thought to be disease-free after initial surgery, was associated with improved DFS.

In contrast, among stage I patients, postoperative RAI appeared to be associated with worse OS (Table 4). Death was reported in seven stage I patients, all of whom had received postoperative RAI. However, further evaluation of these fatalities failed to suggest linkage to a potential adverse effect from RAI itself. Only one patient's death was attributed to thyroid cancer: a 35-year-old man with HCC who

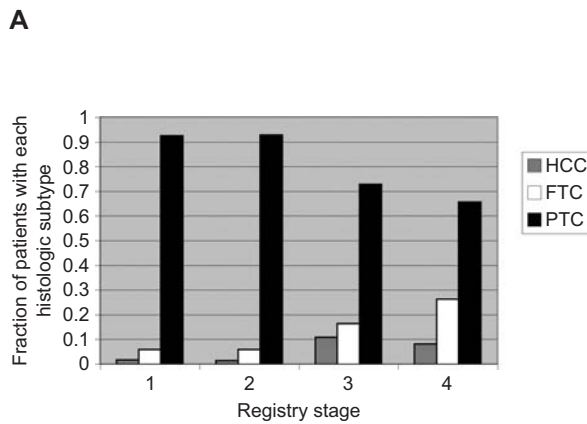


FIG. 1A. Fraction of each histologic subtype within patients in each registry stage. HCC = Hürthle cell variant; FTC = follicular thyroid cancer; PTC = papillary thyroid cancer.

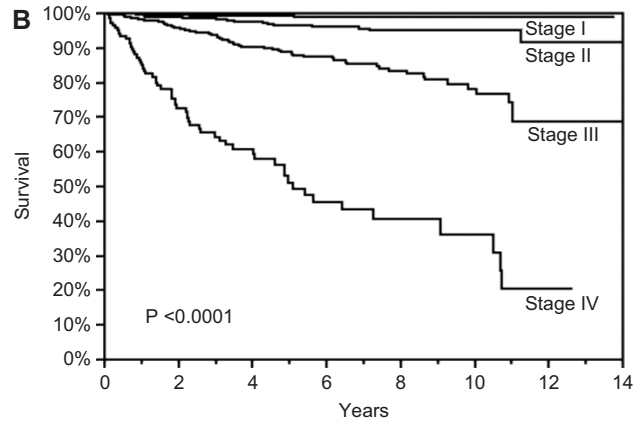


FIG. 1B. Product-limit estimates of overall survival after diagnosis of differentiated thyroid cancer (DTC) according to registry stage assigned at entry. $P < 0.0001$

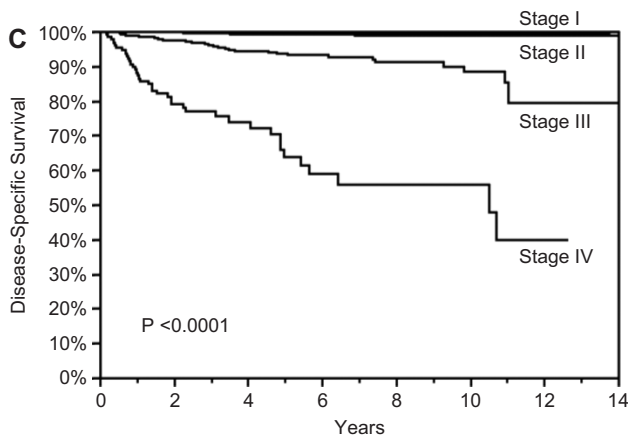


FIG. 1C. Product-limit estimates of disease-specific survival after diagnosis of differentiated thyroid cancer (DTC) according to registry stage assigned at entry. $P < 0.0001$

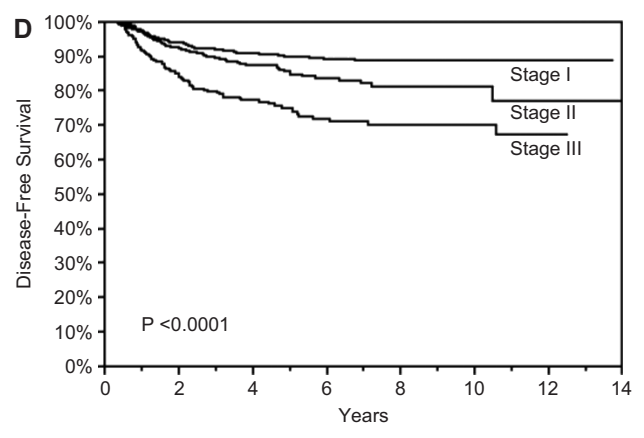


FIG. 1D. Product-limit estimates of disease-free survival after diagnosis of differentiated thyroid cancer (DTC) according to registry stage assigned at entry. $P < 0.0001$

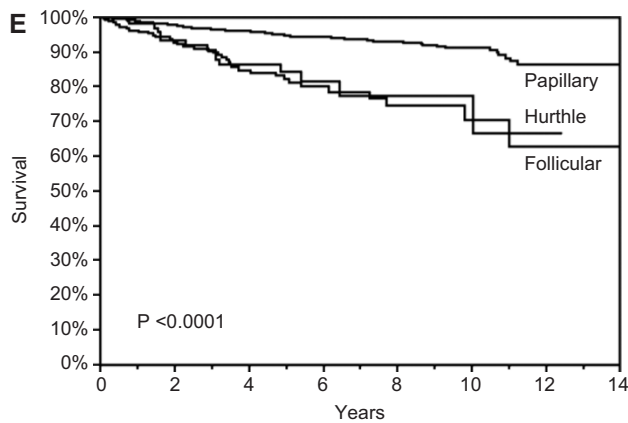


FIG. 1E. Product-limit estimates of overall survival after diagnosis of differentiated thyroid cancer (DTC) according to tumor type. $P < 0.0001$

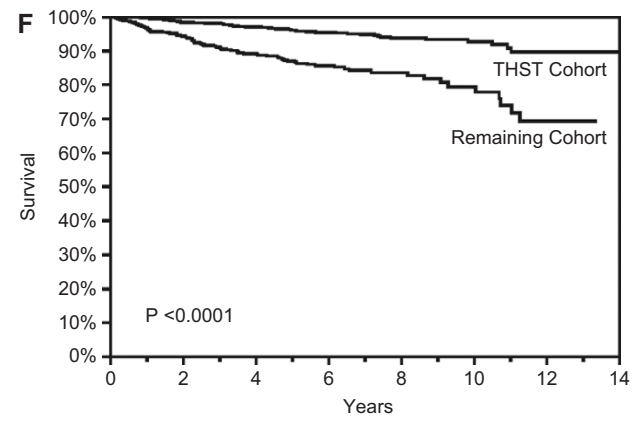


FIG. 1F. Product-limit estimates of overall survival after diagnosis of differentiated thyroid cancer (DTC) according to analysis group (thyroid hormone suppression therapy (THST) cohort versus remaining cohort). $P < 0.0001$

TABLE 3. OUTCOMES FOLLOWING SURGERY: UNIVARIATE ANALYSES, OVERALL COHORT

	Overall Survival			Disease-Specific Survival			Disease-Free Survival		
	RR	95% CI	p	RR	95% CI	p	RR	95% CI	p
Stage I	0.84	0.19–2.04	0.74	0.0013	*	0.28	0.93	0.67–1.25	0.66
Stage II	1.76	1.04–2.83	0.037	0.0013	*	0.35	1.11	0.75–1.54	0.57
Stages III & IV	1.36	1.10–1.66	0.0049	1.3	0.97–1.68	0.073	1.14	0.85–1.48	0.38

RR = risk ratio for outcome, less extensive surgery/NTT = near total thyroidectomy. RR > 1 indicates better outcome associated with NTT. * = lower limit of 95% CI approaches 0. Registry staging is used.

TABLE 4. OUTCOMES FOLLOWING RADIOIODINE: UNIVARIATE ANALYSES, OVERALL COHORT

	Overall Survival			Disease-Specific Survival			Disease-Free Survival		
	RR	95% CI	p	RR	95% CI	p	RR	95% CI	p
Stage I	0.0006	*	0.013	0.00063	*	0.1	0.64	0.47–0.85	0.0013
Stage II	1.71	1.07–2.74	0.026	1.21	0.26–3.92	0.76	1.03	0.75–1.39	0.84
Stages III & IV	1.43	1.17–1.72	0.0006	1.46	1.13–1.87	0.0045	1.32	1.02–1.68	0.035

RR = risk ratio for outcome, no RAI/RAI = radioactive iodine-131. RR > 1 indicates a better outcome associated with RAI. * = lower limit of 95% CI approaches 0. Registry staging is used.

had microscopic extrathyroidal invasion, but in whom extracervical disease was not diagnosed at initial presentation. However, lung metastases were identified 10 months after his initial surgery, and he died 18 months later. Three patients died of other malignancies (breast carcinoma in two, systemic mastocytosis in one), but in these cases the non-thyroid cancer had been diagnosed prior to the diagnosis and treatment of DTC. One patient died of sudden cardiac arrest and another from homicide, both reportedly without evidence of DTC at the time of demise. The cause of death was uncertain in one patient, although she was found to have a superior vena cava thrombus five years after surgery for DTC and 14 months before her last disease-free follow-up examination. Propensity analysis did not alter the conclusions regarding the association between RAI and OS or DSS.

In addition, DFS also appeared to be worse by univariate analysis among stage I patients treated with adjuvant RAI (Table 4). Five- and ten-year DFS rates were 94% and 92%, respectively, for stage I patients not treated with adjuvant RAI, compared with 87% and 86% for those treated. Median time to detection of recurrence was 2.3 years versus 1.2 years for the non-RAI and RAI groups, respectively ($p=0.097$), suggesting earlier detection of recurrence associated with

prior use of adjuvant RAI. Due to the potential impact of bias in selection of patients for adjuvant therapy in this observational study, propensity scores were used to balance covariates that might have influenced DFS (Table 5). Within each propensity score stratum of patients with similar likelihood of having received adjuvant RAI, DFS did not differ significantly between the no RAI and RAI treatment groups. By proportional hazards modeling, the propensity-score adjusted hazard ratio for no RAI compared with adjuvant RAI was no longer a significant predictor of DFS (RR 0.8, 95% CI 0.56–1.10).

Effect of surgery and RAI in the THST cohort

Data regarding TSH levels were insufficient in 47% of registry patients. Therefore, the effects of therapies on outcomes were also analyzed for the THST cohort, in which more than 50% of TSH values were available. In this subgroup, the relative risks associated with surgical extent and postoperative RAI were of similar magnitude and within the 95% confidence intervals of the corresponding analyses in the overall cohort. However, with only 41% of the overall deaths occurring in the THST cohort, the only statistically

TABLE 5. PROPENSITY SCORE ANALYSIS OF RAI (RADIOACTIVE IODINE-131) THERAPY FOR REGISTRY STAGE I PATIENTS, OVERALL COHORT

Propensity Stratum for adjuvant RAI (1 = lowest likelihood, 5 = highest likelihood)	Adjuvant RAI			No adjuvant RAI			RR	95% CI	p
	N	% of stratum	5 yr DFS, %	N	% of stratum	5 yr DFS, %			
1	35	18%	92%	164	82%	95%	0.61	0.25–1.67	0.3
2	88	44%	86%	111	56%	94%	0.66	0.37–1.15	0.15
3	175	77%	92%	53	23%	88%	1.22	0.63–2.18	0.53
4	124	85%	89%	22	15%	100%	0.0008	*	0.12
5	167	88%	78%	23	12%	86%	0.83	0.33–1.52	0.59

RR = risk ratio for outcome, no adjuvant RAI/adjuvant RAI. RR > 1 indicates a better outcome associated with adjuvant RAI. * = lower limit of 95% CI approaches 0.

significant results were improved OS and DSS in high-risk patients associated with RAI treatment (OS: RR 1.74, 95% CI 1.27–2.31; DSS: RR 1.86, 95% CI 1.21–2.77). Patients with lower TSH scores (corresponding to greater degrees of suppression) were more likely to have received postoperative RAI therapy (81% of patients with scores 1.0–1.9, 72% with scores 2.0–2.9, and 64% with scores 3.0–4.0, $p < 0.0001$).

Effect of THST in the THST cohort

THST was associated with improved OS in stage II and high-risk patients in univariate analyses, but did not appear to affect outcomes in stage I patients (Table 6). For the stage II patients, there was a significant association between improved OS and having a mean TSH score <3 , corresponding to TSH levels averaging in the subnormal to undetectable range. However, no further enhancement of OS was identified within the subgroup of patients whose mean TSH scores were <2 , suggesting that there was no incremental benefit from suppressing TSH towards undetectable levels (Fig. 2a). In contrast, high-risk patients with mean TSH scores >2 had a diminished OS compared with those with scores <2 , suggesting that suppression of TSH levels to the limits of assay detection was optimal (Fig. 2b). THST was also associated with improved DSS in high-risk patients, but no effect on DFS was identified for any stage (Table 6).

Multivariate analyses of the effect of surgery, RAI, and THST

Treatments significantly associated with outcomes by univariate methods were entered into multivariate analysis models to identify independent predictors and provide adjusted risk ratios (Table 7, Fig. 3). For high-risk patients in the overall cohort, improved OS was significantly associated with either NTT or RAI therapy. A combination of both these

modalities was more strongly associated with improved OS. In stage II patients, the model that combined both NTT and RAI was significant for improved OS, although neither parameter was independently significant. OS appeared to be worse in stage I patients who had been treated with RAI.

In the overall cohort, improved DSS and DFS were associated with RAI use in high-risk patients; the addition of other treatment modalities did not affect the significant association between RAI and improvement in these outcomes. No effects of treatment on DSS were apparent in stage I and II patients. As previously discussed, a decreased DFS was associated with adjuvant RAI in stage I patients, which was no longer significant after propensity analysis (Table 5). Within the smaller THST cohort, OS was improved in stage II patients with mean TSH scores <3 , and was also improved in high risk patients who had either RAI therapy or mean TSH scores <2 . RAI therapy was associated with improved DSS in high-risk patients. Further improvement in OS and DSS was associated with combined RAI and THST in the high-risk group. There was no association between THST and DFS.

Discussion

Our analysis of registry data provides information regarding clinical outcomes following initial treatment in a large prospectively followed cohort. This information is particularly valuable because more ideal data from randomized controlled trials are unavailable. There are acknowledged limitations in analysis of a registry database. Possible differences in case reporting to the registry, and the impact of the diversity of management practices within participating institutions, are likely to be mitigated by the fact that no institution has contributed more than 15% of the registry patients. Our median follow-up duration is relatively short and the average number of TSH values per patient is limited.

TABLE 6. OUTCOMES ASSOCIATED WITH MEAN TSH SCORES: UNIVARIATE ANALYSES

	Overall Survival				Disease-Specific Survival			Disease-Free Survival		
	TSH score group	RR	95% CI	p^{**}	RR	95% CI	p^{**}	RR	95% CI	p^{**}
Stage I	3/2	5×10^{-6}	*	0.9	1.00	*	1.00	0.53	0.085–1.83	0.38
	2/1	1.16	0.05–29.4		1.00	*		0.79	0.44–1.43	
Stage II	3/2	1×10^8	*	0.0003	3.7×10^{-7}	*	0.61	0.98	0.16–3.49	0.97
	2/1	5×10^{-7}	*		1.00	*		1.09	0.54–2.18	
Stages III & IV	3/2	1.92	0.71–4.44	0.016	3.13	0.98–8.65	0.024	0.76	0.23–1.93	0.56
	2/1	1.88	1.06–3.41		1.76	0.77–4.13		1.31	0.78–2.21	

	Overall Survival				Disease-Specific Survival			Disease-Free Survival		
	TSH score group	RR	95% CI	p^{**}	RR	95% CI	p^{**}	RR	95% CI	p^{**}
Stage I	3/(1&2)	6×10^{-6}	*	0.65	1.00	*	1.00	0.47	0.077–1.54	0.24
Stage II	3/(1&2)	101	10–2300	0.0001	1.6×10^{-6}	*	0.89	1.03	0.17–3.43	0.97
Stages III & IV	(3&2)/1	2.04	1.18–3.64	0.011	4.07	1.32–10.5	0.017	1.27	0.77–2.11	0.35

Mean TSH score groups: group 1 (mean TSH scores*** 1.0–1.9), group 2 (mean TSH scores 2.0–2.9), and group 3 (mean TSH scores 3.0–4.0). ***TSH score = mean of all TSH results where: 1 = undetectable, 2 = subnormal, 3 = normal, 4 = elevated. ** = p value calculated for trend.

THST = thyroid hormone suppression therapy cohort. RR = risk ratio for outcome, higher mean TSH score/lower mean TSH score. RR >1 indicates a better outcome associated with the lower mean TSH score. * = lower limit of 95% CI approaches 0. Registry staging is used.

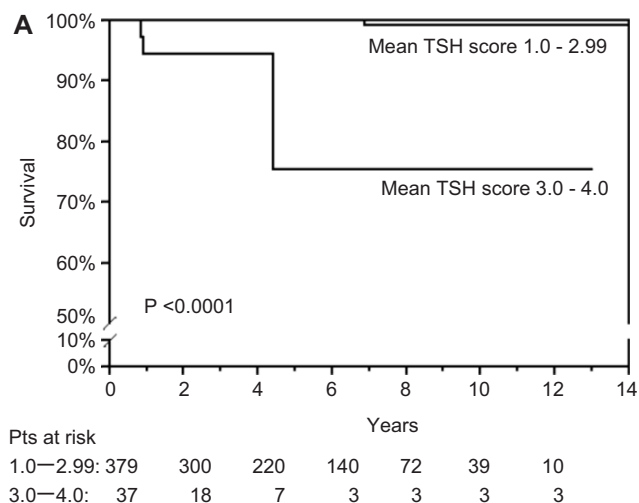


FIG. 2A. Product-limit estimates of overall survival after diagnosis of registry stage II differentiated thyroid cancer (DTC) according to mean TSH score.

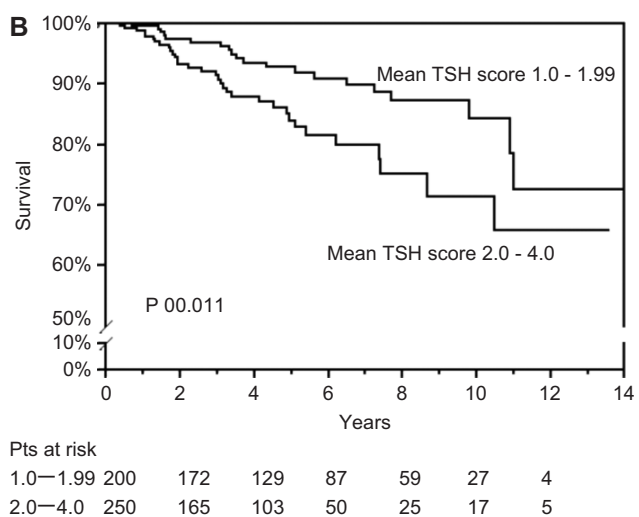


FIG. 2B. Product-limit estimates of overall survival after diagnosis of high-risk differentiated thyroid cancer (DTC) according to mean TSH score.

These deficiencies will presumably be remedied in future registry analyses. Another limitation of our analysis may be that all histologic classifications of PTC and FTC have been merged into one group, as we were unable to identify differences in treatment effects between the histologies. The common management strategies for PTC and FTC are reflected in the recent management guidelines for DTC published by the American Thyroid Association (45). Although the biologic behavior and molecular characteristics of these tumors differ, the poor consensus observed among expert pathologists in the diagnosis of follicular malignancies, the changing patterns of diagnosing follicular tumors in the past 15 years, and the interinstitutional variations in the routine performance of centralized pathology review provide sufficient uncertainty (46–50) to justify combining all DTC tumors in this registry analysis.

Our findings of significantly diminished OS, DSS, and DFS for each successive disease stage (Fig. 1b, 1c, 1d) buttress the previously reported (3) reliability of our staging system. As long as the limitations of extrapolating from a population to an individual are recognized, these data should help provide risk estimates of not only death, but also recurrence, in newly diagnosed DTC patients, thus assisting in the selection of rationale therapies.

Previous studies suggest NTT and RAI are of benefit for high-risk patients. Our current analysis confirms that improved OS is associated with NTT and RAI in high-risk patients, but additionally shows that both these treatments are associated with improved OS in stage II patients also (Tables 3, 4, 7; Fig. 3). In the case of high-risk patients NTT and RAI are associated with improved OS as separate or combined treatments. In contrast, only combined treatment is associated with improved OS in stage II patients. This newly reported benefit of these therapies in low-risk, stage II patients is likely revealed due to our large patient numbers and more extended follow-up.

Within the smaller THST cohort, NTT was no longer significantly associated with improved OS in the high-risk group. Postoperative RAI treatment, however, remained associated with improved OS in high-risk patients (Table 7, Fig. 3). Although RAI combined with NTT was associated with improved OS in stage II patients in the overall cohort, this could not be confirmed in the smaller THST cohort (Table 7, Fig. 3). The disparate effect of NTT and RAI in the overall and THST cohorts could have a number of explanations. There are considerably more recorded deaths within the overall cohort than the THST cohort, leading to greater statistical power. Because the THST cohort was defined on the basis of the frequency of recorded TSH values, bias may have been introduced into this subgroup analysis. For example, the THST cohort may have experienced more intensive treatment and monitoring, which could have included more aggressive THST that may have confounded the analysis of NTT and RAI. This possibility is supported by the observation that survival of the THST cohort was better than the remaining cohort (Fig. 1f). There was no difference in survival between the overall and remaining cohorts.

The benefit of each treatment may also be obscured by the high concordance of patients who had NTT, RAI, and THST. Independent benefits of these three modalities may be hard to demonstrate, even by multivariate analysis, if their impact is achieved through a biologically plausible “common pathway,” involving physiological iodine uptake and TSH responsiveness. Nonetheless, as nearly all high-risk patients in the THST cohort had undergone prior NTT and RAI, the combination of NTT and RAI followed by THST is likely to be required to provide the best outcomes for high-risk patients.

A previous registry analysis (38) showed a trend for THST to lessen disease progression in high-risk PTC patients. Our current data considerably extend this observation by identifying that THST was associated with significantly improved OS in stage II, III, and IV DTC patients (Tables 6, 7; Fig. 3). For stage II patients, THST that yielded serum TSH levels in the subnormal range was associated with improved OS, and no additional improvement was associated with further degrees of THST (Fig. 2a). In contrast, within the high-risk cohort, each successive decrease in average TSH levels was

TABLE 7. MULTIVARIATE ANALYSES OF OUTCOMES FOLLOWING INITIAL THERAPIES

Best Models, Overall Cohort	Overall Survival			Disease-Specific Survival			Disease-Free Survival			
	RR	95% CI	p	RR	95% CI	p	RR	95% CI	p	
Stage I	No RAI/RAI	0.0006	*	0.013	0.013	0.013	0.64	0.47–0.85	0.0013	0.0013
Stage II	<NTT/NTT	1.56	0.90–2.57	0.11		Not significant				
	No RAI/RAI	1.54	0.94–2.52	0.083		Not significant			Not significant	
Stages III & IV	<NTT/NTT	1.26	1.01–1.55	0.045	0.0004	0.0004	1.46	1.13–1.87	0.0045	0.0045
	No RAI/RAI	1.35	1.10–1.64	0.0049		0.0045	1.32	1.02–1.68	0.035	0.035

Best Model, THST Cohort	Overall Survival			Disease-Specific Survival		
	RR	95% CI	p	RR	95% CI	p
Stage II	TSH score 3/1&2	101	10–2300	0.0001	0.0001	0.0001
Stages III & IV	No RAI/RAI	1.65	1.21–2.21	0.0023	1.79	1.16–2.67
	TSH score 3&2/1	1.9	1.07–3.44	0.027	2.02	0.90–4.79

*“P for model” is for the best model employing <NTT/NTT = near-total thyroidectomy, no RAI/RAI = radioactive iodine-131, and higher mean TSH/lower mean TSH score. RR = risk ratio for outcome. RR >1 indicates a better outcome with the more aggressive therapies. * = lower limit of 95% CI approaches 0. Registry staging is used.

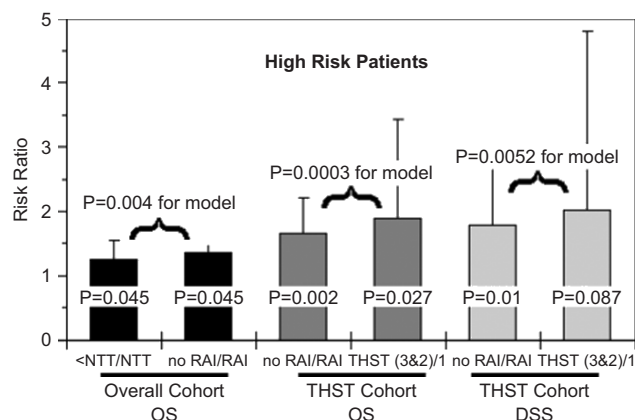


FIG. 3. Multivariate analyses of overall survival (OS) and disease-specific survival (DSS) following initial therapies in high-risk patients. P for model is for the best model employing <NTT/NTT = near total thyroidectomy, no RAI/RAI = radioactive iodine-131, and higher mean TSH/lower mean TSH score. RR = risk ratio for outcome. A RR > 1 suggests a benefit of NTT, RAI, or more aggressive THST = thyroid hormone suppression therapy.

associated with additional improvement in OS, with the highest OS being associated with undetectable to subnormal TSH values (Fig. 2b). Therefore, more aggressive THST is warranted in high-risk patients, whereas less aggressive THST aimed to maintain TSH levels slightly below normal is indicated in low-risk patients.

Despite improved OS associated with NTT in stage II–IV patients, there was only a suggestion of improved DSS in high-risk patients, and no improved DSS in stage II patients with NTT (Table 3). Postoperative RAI treatment, on the other hand was associated with both improved OS and DSS in high-risk patients (Tables 4, 7; Fig. 3). This finding, however, did not extend to stage II patients, in whom only OS was positively associated with RAI. THST was also associated with improved DSS in high-risk patients, but not in stage II patients (Table 6). In the smaller THST cohort, RAI and combined RAI and THST were associated with improved DSS (Table 7, Fig. 3). High-risk patients were found to have an improved DFS associated with RAI. This was the only therapy associated with altered DFS in registry patients (Tables 4, 7). The DFS of stage II patients appeared to be unaltered by RAI therapy.

If treatment were modifying the course of DTC, OS and DSS should be similarly impacted. It is certainly possible that a longer follow-up period might reveal congruent trends in OS and DSS. Possibly, the greater reliability of OS as an endpoint might explain a disparate effect on OS and DSS, as analyses of DSS from registry data may be subject to limited reliability in the attribution of cause of death. Many studies have shown that assignment of the cause of death can be subjective and thus produce biased data (51–53). For example, in the U.K. Trial of Early Detection of Breast Cancer up to 27% of recorded deaths were associated with uncertainty about their cause (54). OS is thus a more robust endpoint than DSS in such studies, although patient age likely has a larger confounding effect on OS. In addition, given that nearly half of

the observed deaths were not attributed to the disease or its treatment, the analysis of DSS is also subject to considerably less statistical power than that for OS. Similar arguments also apply to the degree of reliability of DFS as an endpoint, although the frequency of recurrence is significantly higher and thus the statistical power greater than that seen for analysis of OS or DSS. Thus, it is plausible that an impact on OS may not be associated with a similar impact on DSS, or even DFS, unless collection instruments can be refined, monitoring strategies standardized, and the number of events increased by longer follow-up. The analysis of DFS may also be confounded by the use of RAI, which also enhances subsequent disease detection by thyroglobulin determination and RAI scanning, thus potentially leading to lead-time bias in the early identification of recurrence (26).

Certainly, if only the DSS and DFS results are utilized, a more conservative set of conclusions may be reached. One may alternatively conclude that (i) NTT was not associated with any effect, positive or negative, in any stage patient, (ii) RAI was associated with improved DSS and DFS only in high-risk patients, and (iii) THST was associated with improved DSS only in high-risk patients, but had no impact on DFS. However, if conclusions are based only on the less robust endpoints of DSS and DFS, we believe that the benefits of therapies in DTC patients may be underestimated.

Despite improved outcomes associated with RAI in stages II–IV patients, RAI was not associated with any obvious benefit in stage I patients, and at first appeared associated with worse OS and DFS (Table 4). The apparent association of RAI with worse outcomes in this group is difficult to explain as a radiotherapeutic side effect. Lack of benefit from RAI therapy might be anticipated for the subset of stage I patients who had primary tumors smaller than 1 cm. However, a worse OS would presumably not be expected. It has been postulated that RAI could induce mutations within remnant thyroid tissue and/or cancer (55,56), and that the selective ablation of more differentiated thyroid cells allows subsequent growth of less differentiated clones (57). However, our data do not support this theory, as only one stage I patient died of thyroid cancer. Neither did our data suggest second primary cancers (58) occurring after thyroid cancer were the cause of death in stage I patients, as three patients died due to advanced malignancies that preceded the diagnosis and RAI treatment of their stage I DTC.

The decision to administer RAI appeared to be based on clinical characteristics of concern to the treating physicians, as demonstrated by propensity analysis, and suggests that selection bias may have produced an apparent association with decreased DFS (Table 5). An adverse effect of RAI on DFS, which was not statistically significant, has been reported in low risk patients previously (59), but such reduced DFS could reflect lead-time bias due to greater specificity of testing to detect recurrence after thyroid bed ablation. Although biases due to lead-time, length of preclinical stage, and differential application of detection methods have been discussed with regard to cancer epidemiology (60,61), the application of RAI both to destroy DTC cells and to improve subsequent disease detection may be unique in the oncology literature. Nevertheless, our data suggest that postoperative RAI therapy does not provide significant benefit in stage I patients, and could even be harmful. Further analysis of

stage I patients by examining outcomes in smaller subgroups was not possible due to the small number of events. Should an absence of benefit following RAI in stage I patients be confirmed in additional studies with longer follow-up, this would substantiate a more conservative approach to use of this treatment in stage I patients.

Despite some caveats, we believe that our analyses of this large patient group offer new and valuable information regarding the impact of specific therapies on outcomes in DTC. Our data provide new insight into the management of low-risk patients by demonstrating that NTT and RAI followed by a moderate degree of THST improve OS in stage II patients. In agreement with other investigators, our results also indicate that NTT and RAI are associated with longer OS in high-risk DTC patients. Uniquely, our data show that more aggressive THST, yielding a subnormal or undetectable TSH, is associated with improved OS in high-risk patients, whereas such aggressive therapy does not appear necessary in lower risk patients. Further accumulation of follow-up data may be necessary to extend these observations to DSS and DFS in stage II, III, and IV DTC patients. Improved OS of stage I patients did not appear to be associated with any particular therapy. Further evaluation of the potential risks and benefits of treatment in stage I patients is, therefore, indicated.

Acknowledgments

The NTCTCSG is supported in part by a research grant from Genzyme Corporation, and unrestricted donations from Knoll Pharmaceuticals, Inc., KRONUS, Inc., and Saudi Arabian Oil Company. Jacqueline Jonklaas is supported by a Career Development Award, K23 RR16524, from the National Center for Research Resources. Kenneth B. Ain has a Veterans Affairs Merit Award. The authors, as principal investigators at each NTCTCSG-institution, would like to thank the physicians and staff members who participated in the management and follow-up of these patients. They would like to acknowledge the prodigious contributions of the institutional research staff who collected and submitted the data, including Shehnaz Bana, Marge E. Ewertz, R.N., Beverly McLaughlin, R.M.A., Michele J. Smith, B.S.N., Terry Stauber, R.N., Naomi Walpert, R.N., and Marina Zemskova, M.D. They also appreciate the considerable assistance provided by Matthew Sperling and Roberto Reyes for computer support. Finally, the authors would like to acknowledge the efforts of the numerous physicians and scientists whose contributions were critical to the establishment of the Registry, including Bonnie Specker, John Simpson, Herbert Klein, Ian Hay, Ronald Nishiyama, Leslie DeGroot, William Beierwaltes, Robert Leeper, William McConahey, Katsuro Shimaoka, Terry Taylor, and Irwin Klein.

References

- Jemal A, Murray T, Ward E, *et al.* 2005 Cancer statistics, 2005. *CA Cancer J Clin* **55**:10–30.
- Mazzaferri EL 2006 Managing small thyroid cancers. *JAMA* **295**:2179–2182.
- Sherman SI, Brierley JD, Sperling M, *et al.* 1998 Prospective multicenter study of thyroid carcinoma treatment: initial analysis of staging and outcome. National Thyroid Cancer Treatment Cooperative Study Registry Group. *Cancer* **83**:1012–1021.
- Udelsman R, Lakatos E, Ladenson P 1996 Optimal surgery for papillary thyroid carcinoma. *World J Surg* **20**:88–93.
- Hundahl SA, Cady B, Cunningham MP, *et al.* 2000 Initial results from a prospective cohort study of 5583 cases of thyroid carcinoma treated in the United States during 1996. U.S. and German Thyroid Cancer Study Group. An American College of Surgeons Commission on Cancer Patient Care Evaluation study. *Cancer* **89**:202–217.
- Spencer CA, Takeuchi M, Kazarosyan M, *et al.* 1998 Serum thyroglobulin autoantibodies: prevalence, influence on serum thyroglobulin measurement, and prognostic significance in patients with differentiated thyroid carcinoma. *J Clin Endocrinol Metab* **83**:1121–1127.
- Massin JP, Savoie JC, Garnier H, Guiraudon G, Leger FA, Bacourt F 1984 Pulmonary metastases in differentiated thyroid carcinoma. Study of 58 cases with implications for the primary tumor treatment. *Cancer* **53**:982–992.
- Miccoli P, Antonelli A, Spinelli C, Ferdeghini M, Fallahi P, Baschieri L 1998 Completion total thyroidectomy in children with thyroid cancer secondary to the Chernobyl accident. *Arch Surg* **133**:89–93.
- Wartofsky L, Sherman SI, Gopal J, Schlumberger M, Hay ID 1998 The use of radioactive iodine in patients with papillary and follicular thyroid cancer. *J Clin Endocrinol Metab* **83**:4195–4203.
- DeGroot LJ, Kaplan EL, Straus FH, Shukla MS 1994 Does the method of management of papillary thyroid carcinoma make a difference in outcome? *World J Surg* **18**:123–130.
- Sanders LE, Rossi RL 1995 Occult well differentiated thyroid carcinoma presenting as cervical node disease. *World J Surg* **19**:642–646; discussion 646–647.
- Wanebo H, Coburn M, Teates D, Cole B 1998 Total thyroidectomy does not enhance disease control or survival even in high-risk patients with differentiated thyroid cancer. *Ann Surg* **227**:912–921.
- Haigh PI, Urbach DR, Rotstein LE 2005 Extent of thyroidectomy is not a major determinant of survival in low- or high-risk papillary thyroid cancer. *Ann Surg Oncol* **12**:81–89.
- Mazzaferri EL, Kloos RT 2001 Clinical review 128: current approaches to primary therapy for papillary and follicular thyroid cancer. *J Clin Endocrinol Metab* **86**:1447–1463.
- Mazzaferri EL 1997 Thyroid remnant 131I ablation for papillary and follicular thyroid carcinoma. *Thyroid* **7**:265–271.
- Tsang RW, Brierley JD, Simpson WJ, Panzarella T, Gospodarowicz MK, Sutcliffe SB 1998 The effects of surgery, radioiodine, and external radiation therapy on the clinical outcome of patients with differentiated thyroid carcinoma. *Cancer* **82**:375–388.
- Chow SM, Law SC, Mendenhall WM, *et al.* 2002 Papillary thyroid carcinoma: prognostic factors and the role of radioiodine and external radiotherapy. *Int J Radiat Oncol Biol Phys* **52**:784–795.
- Samaan NA, Schultz PN, Hickey RC, *et al.* 1992 The results of various modalities of treatment of well differentiated thyroid carcinomas: a retrospective review of 1599 patients. *J Clin Endocrinol Metab* **75**:714–720.
- Krishnamurthy GT, Bland WH 1977 Radioiodine I-31 therapy in the management of thyroid cancer. A prospective study. *Cancer* **40**:195–202.
- Petrich T, Widjaja A, Musholt TJ, *et al.* 2001 Outcome after radioiodine therapy in 107 patients with differentiated thy-

- roid carcinoma and initial bone metastases: side-effects and influence of age. *Eur J Nucl Med* **28**:203–208.
21. Lerch H, Schober O, Kuwert T, Saur HB 1997 Survival of differentiated thyroid carcinoma studied in 500 patients. *J Clin Oncol* **15**:2067–2075.
 22. Taylor T, Specker B, Robbins J, *et al.* 1998 Outcome after treatment of high-risk papillary and non-Hurthle-cell follicular thyroid carcinoma. *Ann Intern Med* **129**:622–627.
 23. Shoup M, Stojadinovic A, Nissan A, *et al.* 2003 Prognostic indicators of outcomes in patients with distant metastases from differentiated thyroid carcinoma. *J Am Coll Surg* **197**:191–197.
 24. Sawka AM, Thephamongkhon K, Brouwers M, Thabane L, Browman G, Gerstein HC 2004 Clinical review 170: A systematic review and metaanalysis of the effectiveness of radioactive iodine remnant ablation for well-differentiated thyroid cancer. *J Clin Endocrinol Metab* **89**:3668–3676.
 25. Durante C, Haddy N, Baudin E, *et al.* 2006 Long term outcome of 444 patients with distant metastases from papillary and follicular thyroid carcinoma: benefits and limits of radioiodine therapy.
 26. Hay ID, Thompson GB, Grant CS, *et al.* 2002 Papillary thyroid carcinoma managed at the Mayo Clinic during six decades (1940–1999): temporal trends in initial therapy and long-term outcome in 2444 consecutively treated patients. *World J Surg* **26**:879–885.
 27. Derwahl M, Broecker M, Kraiem Z 1999 Clinical review 101: thyrotropin may not be the dominant growth factor in benign and malignant thyroid tumors. *J Clin Endocrinol Metab* **84**:829–834.
 28. Hoelting T, Tezelman S, Siperstein AE, Duh QY, Clark OH 1995 Biphasic effects of thyrotropin on invasion and growth of papillary and follicular thyroid cancer *in vitro*. *Thyroid* **5**:35–40.
 29. Tanaka K, Fukunari N, Nagahama M, Sugino K, Ito K, Shimizu K 2003 Long term effect of TSH suppression therapy on papillary carcinoma with distant metastases. 75th Annual Meeting of the American Thyroid Association. Abstract 132.
 30. Balme H 1954 Metastatic carcinoma of the thyroid successfully treated with thyroxine. *Lancet* **1**:812–813.
 31. Mazzaferri EL, Jhiang SM 1994 Long-term impact of initial surgical and medical therapy on papillary and follicular thyroid cancer. *Am J Med* **97**:418–428.
 32. Staunton MD 1994 Thyroid cancer: a multivariate analysis on influence of treatment on long-term survival. *Eur J Surg Oncol* **20**:613–621.
 33. Young RL, Mazzaferri EL, Rahe AJ, Dorfman SG 1980 Pure follicular thyroid carcinoma: impact of therapy in 214 patients. *J Nucl Med* **21**:733–737.
 34. Simpson WJ, Panzarella T, Carruthers JS, Gospodarowicz MK, Sutcliffe SB 1988 Papillary and follicular thyroid cancer: impact of treatment in 1578 patients. *Int J Radiat Oncol Biol Phys* **14**:1063–1075.
 35. Cunningham MP, Duda RB, Recant W, Chmiel JS, Sylvester JA, Fremgen A 1990 Survival discriminants for differentiated thyroid cancer. *Am J Surg* **160**:344–347.
 36. Rossi RL, Cady B, Silverman ML, *et al.* 1988 Surgically incurable well-differentiated thyroid carcinoma. Prognostic factors and results of therapy. *Arch Surg* **123**:569–574.
 37. Sanders LE, Cady B 1998 Differentiated thyroid cancer: re-examination of risk groups and outcome of treatment. *Arch Surg* **133**:419–425.
 38. Cooper DS, Specker B, Ho M, *et al.* 1998 Thyrotropin suppression and disease progression in patients with differentiated thyroid cancer: results from the National Thyroid Cancer Treatment Cooperative Registry. *Thyroid* **8**:737–744.
 39. Pujol P, Daures JP, Nsakala N, Baldet L, Bringer J, Jaffiol C 1996 Degree of thyrotropin suppression as a prognostic determinant in differentiated thyroid cancer. *J Clin Endocrinol Metab* **81**:4318–4323.
 40. McGriff NJ, Csako G, Gourgiotis L, Lori CG, Pucino F, Sarlis NJ 2002 Effects of thyroid hormone suppression therapy on adverse clinical outcomes in thyroid cancer. *Ann Med* **34**:554–564.
 41. Biondi B, Filetti S, Schlumberger M 2005 Thyroid-hormone therapy and thyroid cancer: a reassessment. *Nat Clin Prac Endocrinol Metab* **1**:32–38.
 42. Fleming I, Cooper J, Henson D, Hutter R, Kennedy B, Murphy G (eds) 1997 American Joint Committee on Cancer Staging Manual. 5th ed. Lippincott-Raven, Philadelphia.
 43. D'Agostino RB, Jr 1998 Propensity score methods for bias reduction in the comparison of a treatment to a non-randomized control group. *Stat Med* **17**:2265–2281.
 44. Rubin DB 1997 Estimating causal effects from large data sets using propensity scores. *Ann Intern Med* **127**:757–763.
 45. Cooper DS, Doherty GM, Haugen BR, *et al.* 2006 Management guidelines for patients with thyroid nodules and differentiated thyroid cancer. *Thyroid* **16**:109–142.
 46. Franc B, Valenty M, Galakhin K, *et al.* 2003 Histological validation of diagnoses of thyroid cancer among adults in the registries of Belarus and the Ukraine. *Br J Cancer* **89**:2098–2103.
 47. Hirokawa M, Carney JA, Goellner JR, *et al.* 2002 Observer variation of encapsulated follicular lesions of the thyroid gland. *Am J Surg Pathol* **26**:1508–1514.
 48. Fassina AS, Montesco MC, Ninfo V, Denti P, Masarotto G 1993 Histological evaluation of thyroid carcinomas: reproducibility of the "WHO" classification. *Tumori* **79**:314–320.
 49. Lloyd RV, Erickson LA, Casey MB, *et al.* 2004 Observer variation in the diagnosis of follicular variant of papillary thyroid carcinoma. *Am J Surg Pathol* **28**:1336–1340.
 50. LiVolsi VA, Asa SL 1994 The demise of follicular carcinoma of the thyroid gland. *Thyroid* **4**:233–236.
 51. Goldacre MJ 1993 Cause-specific mortality: understanding uncertain tips of the disease iceberg. *J Epidemiol Community Health* **47**:491–496.
 52. Ravakhah K 2006 Death certificates are not reliable: re-vivification of the autopsy. *South Med J* **99**:728–733.
 53. D'Amico M, Agazzino E, Biagino A, Simonetti A, Marinelli P 1999 Ill-defined and multiple causes on death certificates—a study of misclassification in mortality statistics. *Eur J Epidemiol* **15**:141–148.
 54. Chamberlain J, Coleman D, Ellman R, Moss S 1991 Verification of the cause of death in the trial of early detection of breast cancer. UK Trial of Early Detection of Breast Cancer Group. Trial Co-ordinating Centre. *Br J Cancer* **64**:1151–1156.
 55. Lambert V, Thierens H, Monsieurs M, Roncancio C, Laurent C 2001 Translocation frequencies measured in patients one year after radioactive iodine therapy for thyrotoxicosis. *Int J Radiat Biol* **77**:679–685.
 56. Gutierrez S, Carbonell E, Galofre P, Creus A, Marcos R 1999 Cytogenetic damage after 131-iodine treatment for hyperthyroidism and thyroid cancer. A study using the micronucleus test. *Eur J Nucl Med* **26**:1589–1596.

57. Williams ED 1986 Relation between ¹³¹I therapy for thyrotoxicosis and development of thyroid carcinoma. *Lancet* **2(8504)**:456–457.
58. Sandeep TC, Strachan MW, Reynolds RM, *et al.* 2006 Second primary cancers in thyroid cancer patients: a multinational record linkage study. *J Clin Endocrinol Metab* **91**:1819–1825.
59. Hay ID, McConahey WM, Goellner JR 2002 Managing patients with papillary thyroid carcinoma: insights gained from the Mayo Clinic's experience of treating 2,512 consecutive patients during 1940 through 2000. *Trans Am Clin Climatol Assoc* **113**:241–260.
60. Abel U, Wollermann C 2003 Methodological aspects of the evaluation of postoperative cancer surveillance. Part I: validity. *Clin Lab* **49**:367–377.
61. Prorok PC 1992 Epidemiologic approach for cancer screening. Problems in design and analysis of trials. *Am J Pediatr Hematol Oncol* **14**:117–128.

Address reprint requests to:

Jacqueline Jonklaas
Bldg. D, Ste. 230

Georgetown University Medical Center
4000 Reservoir Rd., N.W.
Washington, DC 20007

E-mail: jj@bc.georgetown.edu